

[0006] Fig.1 is a schematic diagram that shows packets of data of a voice data 20 being sent across a network 10. For example, the data 20 can be media output from the Internet such as a VoIP transmission, a person's voice being transmitted across a wireless phone network, or similar communications data. The data 20 includes audible ranges 20a, 20c, and 20e where there is discernable audio information and silent ranges 20b and 20d where there is an absence of discernable audio information. A sender 12, being a PC, wireless telephone, or other device, sends packets P1-P15 in order at regular intervals, but because of network delay delaying the transmission of the packets P1-P15 some of the packets P1-P15 arriving at a receiver 14, a similar PC or device, must be further delayed by different amounts to form a cohesive voice data 22. The voice data 22 includes audible ranges 22a, 22c, and 22e and silent ranges 22b and 22d corresponding to the ranges 20a-20e of the sent data 20.

[0007] The packet P1 is sent by the sender 12 at a given time. The packet P1 is delayed by the network 10 for any number of reasons, said delay and further delays being indicated in Fig.1 by a shaded block. The packet P1 is further delayed by the receiver 14 so it can be played contiguously with the packet P2 that is also delayed by the network 10. If the packet P1 is not further delayed by the receiver 14, packets P1 and P2 would not be played contiguously, and an audible break in the data 22 would occur. The audible break in the data 22 would be heard by a listener at the receiver 14, which translates to poor audio quality of the playout data 22.

[0008] The packets P2-P5 are all delayed by the network 10 by the same amount of time and do not have to be further delayed by the receiver 14 to be played in sequence with proper timing. However, the packet P7 arrives before the packet P6. The receiver 14 must delay the playout of the packet P7 until the packet P6 is received. This delay is added to the silent range 22b of the data 22 so that the audible range 22c is not affected. The packets P8 and P9 arrive simultaneously as do the packets P10 and P11 because of network delay and packet bursting. Playout of the packets P9 and P11 is accordingly delayed, however, no further delay of the data 22 results. The packets P13 and P14 suffer a similar disorder as the packets P6 and P7. The packets P12 and P15 arrive at the receiver 14 normally.

[0009] The above description with reference to Fig.1 is a simplification. The packets P1-

P15 were assumed to arrive at the receiver delayed by an integer multiple of their packet length. In reality, a substantially large number of packets in a given transmission must be delayed, as network delay and jitter are essentially continuous in time and packet length is digital.

[0010] Fig.1 shows that the entire received data 22 is delayed by three blocks by a combination of network delay and additional playout delay added by the receiver 14. If this additional delay were not added by the receiver 14, some packets would be played out of order and others would not be played at all. The prior art teaches a number of ways to estimate the delay required to be added by the receiver 14. However, too much playout delay can result in inconvenience and even misunderstanding between parties in two-way schemes. A fundamental and arguably most useful method of estimating playout delay is the mean delay and variance (MDV) method described in R. Ramjee, J. Kurose, D. Towsley, and H. Schulzrinne, "Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks", Proceedings of IEEE INFOCOM, Toronto, Canada, pp. 680 – 686, June 1994, which is incorporated herein by reference. The MDV method is further described in Marco Rocchetti, Vittorio Ghini, Giovanni Pau, Paola Salomoni, and Maria Elena Bonfigli, "Design and Experimental Evaluation of an Adaptive Playout Delay Control Mechanism for Packetized Audio for use over the Internet", November 1998, which is also incorporated herein by reference. Briefly, the MDV method estimates playout delay from a variance of a mean network delay in conjunction with a smoothing factor. This simple adaptive approach offers significant improvement over other non-adaptive approaches.

[0011] Another method of estimating playout delay is described in the real-time transport protocol (RTP) standard. H. Schulzrinne, S. Casner, R. Frederick, V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", RFC 1889, January 1996 details the RTP standard and is incorporated herein by reference. The RTP method of estimating delay is essentially the MDV method applied with a fixed smoothing factor. While simpler than the MDV method, the RTP method offers a less accurate estimation of network delay.

[0012] Other prior art methods of estimating playout delay include a spike detection

method described in "Adaptive Playout Mechanisms for Packetized Audio Applications in Wide-Area Networks", and a related gap-based method described in Jesus Pinto and Kenneth J. Christensen, "An Algorithm for Playout of Packet Voice based on Adaptive Adjustment of Talkspurt Silence Periods", 1999. Both the spike detection method and the gap-based method offer little significant improvement over the MDV method at the expense of added complexity.

[0013] Finally, the prior art offers a normalized least mean square (NLMS) method that is described in Phillip DeLeon and Cormac J. Sreenan, "An Adaptive Predictor for Media Playout buffering", Stanford, March 2001. The NLMS is a complicated method that offers no readily apparent advantages over other methods.

[0014] In addition, the prior art has numerous patents relating to the playout of digital information and performance monitoring of the playout. For instance, Daum et al. teach stream synchronization for MPEG playback in the comprehensive US Patent 5,815,634, and Jain describes a real-time receiver and method for receiving and playing out real-time packetized data in US 6,259,677, both of which are included herein by reference. Additionally, Schulman in US 5,600,632 teaches performance monitoring in a network using synchronized network analyzers relating to packet delay, and Agrawal et al. provide a predictive approach to synchronization using a method for maintaining and updating statistical trends of network delay in US 6,072,809, both of which are include herein by reference.

[0015] The prior art methods mentioned and described above optimize the playout delay from network statistics and do not adequately consider a duplex of communications. That is, playout delay is not adjusted considering half-duplex or full-duplex modes of communication. Thus, advantages of considering these modes are not realized resulting in actual playout quality, and therefore overall communications quality, being deteriorated.

Summary of Invention

[0016] It is therefore a primary objective of the claimed invention to provide a duplex aware adaptive playout method and communications device to optimize playout delay of packet communication to solve the problems of the prior art.

[0017] Briefly summarized, the claimed invention method includes detecting a packet communication mode as either a full-duplex mode or a half-duplex mode, calculating a playout delay for a current packet based on the detected packet communication mode, and delaying playout of the current packet by the calculated playout delay. When the packet communication mode is detected to be the half-duplex mode, the calculated playout delay for the current packet is set longer than when the packet communication mode is detected to be the full-duplex mode.

[0018] According to the claimed invention, the method is further capable of calculating a mean network delay variance for the current packet, and determining an estimated jitter for the current packet referencing the mean network delay variance. Then, the playout delay for the current packet is calculated further based on the estimated jitter of the current packet and playout delays of previous packets.

[0019] According to the claimed invention, a communications device includes a playout buffer that receives and buffers packets for playout, a playout controller that controls the playout buffer according to playout delays, a network delay estimator that calculates estimated network delays for the playout controller, and an active detector that detects a packet communication mode for the playout controller. The playout controller determines playout delays of packets from the estimated network delays and the detected packet communication mode. The packet communication mode can be a full-duplex mode or a half-duplex mode.

[0020] It is an advantage of the claimed invention that delaying playout of packets by a longer delay in the half-duplex mode improves playout quality.

[0021] It is an advantage of the claimed invention that delaying playout of packets by a shorter delay in the full-duplex mode accelerates playout response.

[0022] It is a further advantage of the claimed invention that the active detector can detect the communication mode and the playout controller can delay playout of packets according to the communication mode.

[0023] These and other objectives of the claimed invention will no doubt become obvious to those of ordinary skill in the art after reading the following detailed description of the preferred embodiment that is illustrated in the various figures and drawings.

Brief Description of Drawings

[0024] Fig.1 is a schematic diagram showing packets of voice data being sent across a network.

[0025] Fig.2 is a block diagram of a communications device according to the present invention.

[0026] Fig.3 is a flowchart of operation of the communications device according to the present invention.

Detailed Description

[0027] The present invention duplex aware adaptive playout method is best understood when described in conjunction with a communications device. Notation is consistent between all equations and procedures given.

[0028]

Please refer to Fig.2. Fig.2 shows system architecture of a communications device 30 according to the present invention. The communications device 30 comprises a receiver 32 for receiving packets from the network 10, and a playout buffer 34 for receiving packets forwarded by the receiver 32 and for outputting data of the packets to a media output device 36. The playout buffer 34 is used to absorb network jitter so playout of the packets at the media output device 36 is substantially smooth and continuous. The media output device 36 can be a typical media output device such as a voice over Internet protocol (VoIP) player, a speaker or display of a wireless telephone, or other such device. The playout device 30 further comprises a network delay estimator 38 for estimating network delay of the network 10, an active detector 40 for detecting a packet communication mode, and a playout controller 42 for controlling the playout buffer 34. The active detector 40 determines if the packet communication mode is a half-duplex mode, when a media input device 44 is inactive, or is a full-duplex mode, when the media input device 44 is active. The media input device 44, being a device corresponding to the media output device 36, accepts information input to the communications device 30 and forwards packets to a transmitter 46. The transmitter 46 outputs packets to the network 10. The playout controller 42 sets the delays of received packets in the playout buffer 34 according to network delay estimates from the network delay estimator 38 and the packet

communication mode as determined by the active detector 40. Finally, when the playout delay of a packet expires, the playout buffer 34 sends the packet to the media output device 36 for playout.

[0029] The active detector 40 detects if the communications device 30 is in the half-duplex mode or the full-duplex mode. If the media input device 44 is not receiving information, or the amount of information is below a threshold (for example, voice volume is nearly silent), the active detector 40 detects that the communications device 30 is in the half-duplex mode. Conversely, when the media input device 44 is receiving information above the threshold, the full-duplex mode is detected. That is, the half-duplex mode is one-way communication and the full-duplex mode is two-way as is well understood in the art. Other methods of switching between the half-duplex and full-duplex modes are also possible in relation to the present invention:

[0030] The network delay estimator 38 and playout controller 42 estimate network delay and mean network delay variance according to the following equations:

[0031]

$$\begin{aligned} D_i &= |(R_i - R_{i-1}) - (S_i - S_{i-1})| \\ MD_i &= F \times MD_{i-1} + (1 - F) \times D_i \\ V_i &= |MD_i - D_i| \\ MV_i &= F \times MV_{i-1} + (1 - F) \times V_i \end{aligned}$$

(Eqns.1)

[0032] where,

[0033] D is network delay;

[0034] R is a receiver timestamp;

[0035] S is a sender timestamp;

[0036] i is an index that denotes a current packet;

[0037] i-1 is an index that denotes a previous packet;

[0038] MD is mean network delay;

[0039] F is a smoothing factor;

[0040] V is network delay variance;

[0041] MV is mean network delay variance;

[0042] Essentially, the network delay estimator 38 first estimates a network delay for the current packet based on network statistics of the current packet and of the previous packets. The network delay estimator 38 then forwards the estimated network delay for the current packet to the playout controller 42. The playout controller 42 then calculates a mean network delay variance for the current packet using the mean network delay, the smoothing factor, and the network delay variance. The playout controller 42 then calculates a playout delay for the current packet according to the following procedure:

[0043] (Eqns.2)

$$\begin{aligned} \text{NTD}_1 &= \frac{\text{TD}_1}{\text{MAX_TD}} \\ \text{DMOS}_1 &= \text{MAX_MOS} \times (1 - \text{NTD}_1)^{\text{DF}} \end{aligned}$$

[0044] where,

[0045] M is the packet communication mode, wherein FULL_DUPLEX and HALF_DUPLEX are valid values;

[0046] J is an estimated network jitter;

[0047] FSF is a full-duplex scaling factor;

[0048] PD is the playout delay;

[0049] FF is a full-duplex smoothing factor;

[0050] HSF is a half-duplex scaling factor;

[0051] HF is a half-duplex smoothing factor;

[0052] Referencing the procedure described in Eqns.2, the playout controller 42 first determines the mode of packet communication, then determines the network jitter,

and finally calculates the playout delay for the current packet. The determined network jitter depends on the network delay variance of Eqns.1, a scaling factor. The playout delay depends on the jitter measurement for the current packet, a smoothing factor, and the actual playout delays for the previous packets. The playout controller 42 selects the scaling and smoothing factors depending on the mode of packet communication, either FULL_DUPLEX or HALF_DUPLEX, as determined by the active detector 40.

[0053] For optimum performance, the half-duplex scaling factor HSF is set larger than the full-duplex scaling factor FSF. The reason being a longer playout delay is permissible in half-duplex mode, and is conducive to higher playout quality. For example, if one person is talking and another person is listening (i.e. for one communications device 30 the media input device 44 is inactive) increased playout delay simply results in improved playout quality for the listener. On the contrary, in full-duplex mode, playout delay is minimized (low FSF value relative to HSF) as much as practically possible without unduly sacrificing playout quality to allow smooth simultaneous communication between more than one communications device 30. In full duplex mode, when both people are speaking and listening at the same time, short playout delay is important for communication effectiveness. In full-duplex mode, playout quality is exchanged for shorter playout delay. More to this point, the full-duplex smoothing factor FF is set larger than the half-duplex smoothing factor HF to provide more rapid adjustment of playout delay required by the communications device 30 in full-duplex mode. This causes the playout delay to depend more on the actual current estimated network jitter and less on the playout delays of the previous packets or playout history, thus giving the full-duplex mode relatively quick response. The values of the scaling factors HSF and FSF and the smoothing factors HF and FF are set by and stored in the playout controller 42.

[0054] In practical application, the playout device 30 and its constituent components can be realized using conventional electronic circuits, integrated circuits, and related software programs. The logic and programming of the playout controller 42 and the network delay estimator 38, and response and thresholds of the active detector 40 can be fine-tuned and designed to suit any relevant communications application. Such applications include VoIP players, videophones, on-line games, and the other real-

time interactive communication devices for the Internet and wireless networks. The previously described process for optimizing playout delays of packets performed by the communications device 30 can be summarized with reference to the flowchart of Fig.3 that is described as follows:

- [0055] Step 100:
- [0056] Start. Two (or more) communications devices 30 are communicating through the network 10;
- [0057] Step 102:
- [0058] The active detector 40 detects the mode of packet communication as full-duplex or half-duplex;
- [0059] Step 104:
- [0060] The network delay estimator 38 and the playout controller 42 determine network jitter for the current packet being received at the receiver 32. Procedures in accordance with Eqns.1 are executed;
- [0061] Step 106:
- [0062] The playout controller 42 calculates the playout delay of the current packet according to the detected mode of step 102 and the estimated network delay of step 104, and forwards the playout delay of the current packet to the playout buffer 34. The procedure of Eqns.2 is performed. The playout delay for the current packet is stored in the playout controller 42 for use in a future execution of this procedure (Fig.3) for a next packet;
- [0063] Step 108:
- [0064] The media output device 36 delays playout of packets in the playout buffer 34 according to the playout delays;
- [0065] Step 110: Is communications complete? If not, return to step 102. If it is, go to step 112;
- [0066] Step 112: End.

[0067] In practical application, the above procedure is performed continuously by each connected communications device 30 in the network 10. As each communications device 30 delays playout of received packets according to the duplex, overall communications quality is improved.

[0068] Those skilled in the art will readily observe that numerous modifications and alterations of the device may be made while retaining the teachings of the invention. Accordingly, the above disclosure should be construed as limited only by the metes and bounds of the appended claims.